

# 植物炭疽病真菌金属硫蛋白的三级结构预测\*

何红珍 吕 瞰 赵南明 张日清 刘进元\*\*

(清华大学生物科学与技术系, 生物膜与膜生物工程国家重点实验室, 北京 100084)

**摘要** 根据实验测定的 I 类金属硫蛋白 (metallothionein, MT) 三级结构的实验数据, 给出该类蛋白质的两种特征结构 (CXC、CXXC 一级结构, 半胱氨酸-金属络合簇三级结构) 的原子间距离约束条件, 然后运用距离几何算法计算出系列可能的构象. 从这些构象中经统计分析筛选出目标函数值显著较小的结构作为所预测蛋白质的三级结构模型. 用已知结构的蓝蟹 MT 对方法进行检验证实其可行性后, 对植物炭疽病真菌金属硫蛋白 CAP3 进行了三级结构预测.

**关键词** 金属硫蛋白 CAP3, 三级结构预测, 距离几何算法

**学科分类号** Q51, Q939.5

植物炭疽病真菌孢子发芽时在芽管顶端长出一种称为附着器的结构, 粘着于寄主表皮, 并于其下长出侵入丝, 侵入寄主表皮细胞. 对于这一过程的形态和生理方面的研究已经比较深入, 然而其分子机理, 特别是孢子发芽和分化过程中寄主植物信号分子的作用机制还不清楚. Hwang 等从被鳄梨表层蜡质诱导的附着器中, 筛选到仅在附着器形成过程中表达的三个 cDNA. 其中一个 cDNA 表达生成的, 由 26 个氨基酸组成的小蛋白质称为 CAP3. CAP3 富含半胱氨酸, 与粗糙麦孢菌的铜诱导 MT 高度同源, 而且金属离子能诱导其表达, 表明在炭疽病真菌孢子由芽管分化成附着器的过程中, CAP3 可能起存储或者调节重金属离子浓度的作用<sup>[1]</sup>. CAP3 是如何行使以上功能的呢? 了解其三级结构信息肯定有助于阐明该问题. 另外, 从大小和氨基酸排列方式上看, CAP3 相当于动物 I 类 MT 的一个半胱氨酸富含区 (CR). 在过去的四十多年里, 人们已采用过几乎所有的结构探测手段对 I 类 MT 的 CR 区进行过研究, 积累了大量的实验资料<sup>[2-8]</sup>. 这些研究成果为同一家族的 MT 蛋白结构预测提供了坚实的实验数据. 本研究采用基于距离几何算法的预测方法<sup>[9]</sup>, 并作了部分修改后, 对 CAP3 蛋白进行了三级结构预测.

## 1 方法构建

距离几何算法<sup>[10,11]</sup>是一种用原子间的相对距离表示三维结构信息的方法. 它根据 Blumenthal 定理对原子间的最大距离矩阵——上限矩阵和最小距离矩阵——下限矩阵进行修饰, 以排除那些违背三

角不等式的约束, 从而缩小原子间距离取值范围. 然后在每一给定的距离范围内随机取出一个数, 经数值处理后得到原子坐标就得到了分子的一个可能构象. 应用这一算法预测蛋白质结构, 首先是构建规范该蛋白质形成天然构象所必要而且充分的约束条件. 根据 MT 的结构特点, 对这类蛋白质进行结构预测有必要设置以下两项约束:

### a. CXC、CXXC 片段原子间距离约束

MT 的氨基酸序列排列成多个 CXC (C 代表 Cys, X 代表除 Cys 以外的其他氨基酸) 和 CXXC 片段. 两种片段各自都具有相似的几何构象<sup>[2]</sup>. 对 PDB 蛋白质结构数据库中登录的 MT 数据进行统计分析也发现 CXC、CXXC 各片段中原子间的距离很保守. 据此我们构建了一项 MT 构象约束条件, 首先分别计算出已知结构的三种动物 MT (人、大鼠、兔 MT) 所有 CXC、CXXC 片段的原子间距离矩阵并求出相应距离矩阵的平均矩阵和标准偏差矩阵, 然后将平均矩阵加、减标准偏差矩阵分别生成 CXC、CXXC 片段原子间距离约束的上限矩阵和下限矩阵.

### b. Cys-金属络合结构原子间距离约束

在 MT 中, 每个二价金属离子同四个硫原子以配位键结合, 形成一个近似正四面体的结构<sup>[4]</sup>. 据此构建另一项 MT 构象约束条件, 即将 Cys-金

\* 国家高技术 863 青年基金, 教育部跨世纪人才基金和国家重大基础研究 (G19998010100) 资助项目.

\*\* 通讯联系人.

Tel: 010-62772243, E-mail: liujy@mail.tsinghua.edu.cn

收稿日期: 2001-07-30, 接受日期: 2001-10-30

属络合结构设置为正四面体. 具体设置参照 Arseniev 等<sup>[5]</sup>在兔肝 MT 核磁共振结构解析中使用的方法, 即将金属离子以假原子的形式共价连接到 Cys 的 S 原子上, 键长 0.26 nm, 假原子与 Cys 的 C $\beta$  距离 0.33 nm, Cys 金属络合结构设置成棱长 0.42 nm 的正四面体.

根据上述两项约束条件作出原子间距离约束的上限和下限距离矩阵后, 采用距离几何算法软件 DIANA<sup>[12]</sup> (Distance Geometry Algorithm for NMR applications) 在 SGI 图形工作站上进行结构计算, 得到随机产生的满足约束条件的  $n$  (可设置) 个结构, 选取目标函数值最小即对距离约束违反最少的结构, 用 SYBYL 软件进行能量优化后得到预测结果.

为检验上述两项约束是否充分, 乃至整个方法是否能预测 MT 的三级结构, 先采用该方法预测了已由 NMR 测出结构的蓝蟹 MT<sup>[8]</sup> 的结构, 并将预测结果与实测结构进行了比较. 结果发现当第二项约束即 Cys 金属络合结构约束权重设置为 5, 第一项约束即 CXC、CXXC 片段约束设置为 1 时, 预测结构与实测结构主链原子之间的根均方偏差 (RMSD) 值最小, 即两结构符合得最好 (图 1). 由图 1 可看出  $\alpha$ 、 $\beta$  两个结构域的预测结构与相应的实测结构在主链走向及折叠方式上都基本一致, 其中  $\alpha$  结构域预测结构与相应的实测结构主链原子之间的根均方偏差 (RMSD) 值为 3.12;  $\beta$  结构域为 3.49. 这说明上述两项约束条件对 MT 能够给出足够的约束, 应用上述方法能够对 MT 的三级结构进行预测.



**Fig. 1 View of the blue crab MT tertiary structure showing backbone atoms**

(a) the prediction structure of  $\alpha$  domain (left), the experimentally determined structure of  $\alpha$  domain (right, PDB code: 1DMC); (b) the prediction structure of  $\beta$  domain (left), the experimentally determined structure of  $\beta$  domain (right, PDB code: 1DMC).

## 2 CAP3 三级结构预测

CAP3 是植物炭疽病真菌的一个含 26 个氨基酸的小 MT 蛋白, 其序列为: MSGCGCASTG-TCHCGKDCTCAGCPHK. 从大小和氨基酸排列方式看, 它都相当于动物 I 类 MT 的一个 CR 区.

### 2.1 约束条件的构建

CAP3 蛋白中共有 7 个 Cys, 分别分布于氨基酸序列的第 4、6、12、14、18、20 和 23 位上, 从而形成 3 个 CXC 和一个 CXXC 片段, 对于每一个片段, 按上述方法中第一项约束给出原子间距离上、下限矩阵.

在已知结构的 MT 蛋白 CR 区中, 所有的 Cys 都参与同金属离子的共价络合, 连接一个金属离子的 Cys 称为单联 Cys, 连接两个的称为桥联 Cys; 每一个金属离子同四个 Cys 的 S 原子以配位键结合, 形成一个四面体结构. 假设某一 MT 有  $x$  个 Cys, 其中  $m$  个单联、 $n$  个桥联, 结合  $y$  个金属离子, 则有:  $m + n = x$ ,  $m + 2n = 4y$ ; 方程组取正整数解. 当  $x$  为 7 时, 可有两组解:  $m = 6, n = 1, y = 2$ ;  $m = 2, n = 5, y = 3$ . 根据目前已知结构的 4 种 MT, 含 9 个 Cys 的 5 个结构域和含 11 个 Cys 的 3 个结构域都是单联 Cys 多于桥联 Cys 的实验结果, 我们推测 CAP3 取第一组解, 即 7 个 Cys 中 6 个为单联, 一个为桥联, 共结合 2 个金属离子. 此时, 按自由组合将有 70 种 Cys 金属连接组合模式. 对于每一种组合, 按方法所述分别设置第二项约束即 Cys 金属正四面体络合结构原子间距离约束.

### 2.2 结构计算

对于每一种组合模式构建好上述两项原子间距离约束矩阵 (CXC、CXXC 片段结构约束在所有的组合中是相同的), 并将第一项约束权重设置为 1, 第二项约束权重设置为 5 后, 采用 DIANA 软件在 SGI 图形工作站上进行结构计算. 每种组合生成 50 个构象.

### 2.3 组合模式筛选

70 种组合模式中的每一种生成的 50 个构象的目标函数值的平均值见图 2, 其中最小的是第 54 种组合的平均值 49.20, 经平均值显著性差异  $t$  检验, 该数值与其他平均值有显著性差异 ( $P < 0.05$ ). 因为结构的目标函数值越小, 说明它对约束条件满足程度越高, 第 54 种组合形成的 50 个结构的目标函数值平均值显著小于其他组合的平均

值, 这从总体上反映出它相对其他组合更容易形成满足约束条件的结构. 所以我们选取这种组合模式, 即 4、6、20、23 位 Cys 结合一个金属离子, 12、14、18 和 20 位 Cys 结合另一个金属离子, 其中第 20 位是桥联 Cys.

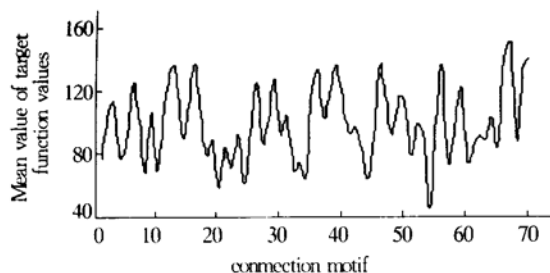


Fig. 2 Mean values of target function values for the 70 kinds of metal-Cys connection motif

#### 2.4 结构优化

将挑选出的第 54 种组合模式再进行结构计算, 生成  $n$  个结构, 选取目标函数值最小的结构.  $n$  由选取的结构中 Cys-金属络合结构决定, 络合结构与正四面体的根均方差值小于 0.3 (参照已知结构的 4 种 MT-1 络合结构数据设置) 即可. 本项计算中  $n$  为 300. 然后将选取的结构进行能量优化, 先采用最速下降法以 0.1 的步长优化 500 轮, 再采用共轭梯度法以 0.05 的步长优化 500 轮. 最后得到的预测结构见图 3. 与动物 MT 的三级结构类似, CAP3 的肽链也折叠成扁盘状, Cys-金属络合簇结构位于盘中间. 肽链上的 7 个 Cys 都参与配位结合金属离子, 共结合 2 个金属离子.

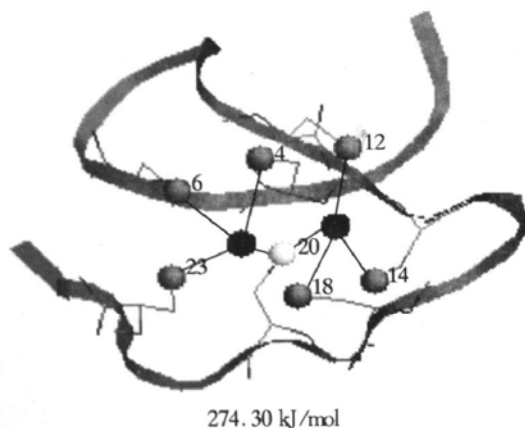


Fig. 3 The prediction structures of CAP3

The bridging sulfurs, terminal sulfurs and metal ions are represented by white, grey and black balls respectively. The conformation energy of the structure is listed below the picture.

### 3 讨论

预测出的 CAP3 结构中, 位于序列中间的 4 个 Cys 结合一个金属离子, 靠近 N 端的 2 个 Cys 和靠近 C 端的 2 个 Cys 结合另一个金属离子, 这从蛋白质折叠的角度来看是可行的. 序列中较均匀分布的 5 个 Gly 更使得这种折叠容易形成. 另外, Messerle 等<sup>[7]</sup>曾作过人 MT-L, 兔 MT-2a 和大鼠 MT-2B 结构域的氨基酸序列比较 (图 4), 发现除兔 MT-2a 多一个氨基酸 A' 外, 三种蛋白质的 9 个 Cys 位于相同的氨基酸序位上. 将 CAP3 的氨基酸序列与前三者比较, 发现只要将 CAP3 氨基酸序位向后移一位, 或者在 N 端添加一个氨基酸, 则其 7 个 Cys 与前三者蛋白质的 9 个 Cys 中的前 7 个的位置完全相同. 根据前三者蛋白质 Cys-金属连接组合模式完全相同的实验结果, 以及在蛋白质进化过程中, 三级结构比一级结构更保守的一般规律, CAP3 中 Cys-金属连接组合模式应该与三种动物 MT 比较类似. 将预测出的 CAP3 中的组合模式与 3 种动物蛋白质中的组合模式进行比较, 发现它们确实很相似: 都是以 4、6、20、23 位上的 Cys 与同一个金属离子配位结合, 12 和 14 位上的 Cys 与另一个金属离子结合. 此外, CAP3 的每个 CXC 片段上的两个 Cys 都与同一个金属离子结合的预测结果也符合 Robbins 等<sup>[3]</sup>在大鼠 MT 结构中观察到的相应的立体化学限制. 另外, 预测结构中 Cys-金属络合结构是非常近似的正四面体 (与正四面体的根均方差值为 0.28), 这是符合该类蛋白质实验结果的, 并且在理论上也是合理的, 因为序列中较均匀分布的 5 个 Gly 使该蛋白质有很大的折叠柔性以使其能形成能量较低的正四面体结构.

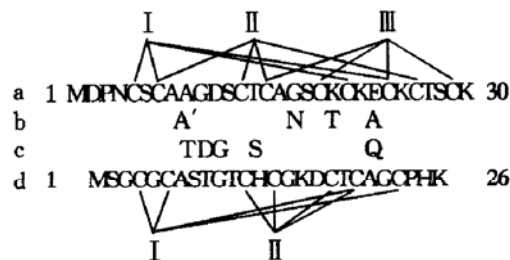


Fig. 4 Comparison of positions of Cys and metal-Cys connection motif of 4 MTs

a, b, c and d represent human MT-2, rabbit MT-2a, rat MT-2 and CAP3 respectively.

如前所述, CAP3 与三种哺乳动物 MT 在结构 (一级结构和三级结构) 上的这些类似性提示了它

们在进化上可能相关联以及在生理功能上可能有相似。因此, CAP3 可能与哺乳动物 MT 类似, 主要以金属离子供体及重金属解毒的作用机制促进炭疽病真菌孢子由芽管分化成附着器。

### 参 考 文 献

- Hwang C S, Kolattukudy P E. Isolation and characterization of genes expressed uniquely during appressorium formation by *Colletotrichum gloeosporioides* conidia induced by the host surface wax. *Mol Gen Genet*, 1995, **247** (3): 282~ 294
- Furey W F, Robbins A H, Clancy L L, *et al.* Crystal structure of Cd, Zn Metallothionein. *Science*, 1986, **231** (4739): 704~ 710
- Robbins A H, McRee D E, Williamson M, *et al.* Refined crystal structure of Cd, Zn metallothionein at 2.0 Å resolution. *J Mol Biol*, 1991, **221** (4): 1269~ 1293
- Braun W, Wagner G, Worgotter E, *et al.* Polypeptide fold in the two metal clusters of metallothionein 2 by NMR in solution. *J Mol Biol*, 1986, **187** (1): 125~ 129
- Arseniev A, Schultze P, Worgotter E, *et al.* Three dimensional structure of rabbit liver [Cd7] metallothionein 2a in aqueous solution determined by nuclear magnetic resonance. *J Mol Biol*, 1988, **201** (3): 637~ 657
- Schultze P, Worgotter E, Braun W, *et al.* Conformation of [Cd7] -metallothionein 2 from rat liver in aqueous solution determined by nuclear magnetic resonance spectroscopy. *J Mol Biol*, 1988, **203** (1): 251~ 268
- Braun W, Messerle B A, Schaeffer A, *et al.* The three-dimensional structure of human metallothionein 2 in solution determined by nuclear magnetic resonance spectroscopy. *J Mol Biol*, 1990, **214** (3): 765~ 779
- Narula S S, Armitage I M, Brouwer M, *et al.* Establishment of two distinct protein domains in blue crab *Callinectes sapidus* metallothionein - I through heteronuclear (1H-113Cd) and homonuclear (1H-1H) correlation NMR experiments. *Magnetic Resonance in Chemistry*, 1993, **31** (Special Issue): S96~ S103
- 朱春明, 吕 曦, 张日清, 等. 猕猴桃金属硫蛋白的结构建模. *科学通报*, 2000, **45** (6): 602~ 607  
Zhu C M, Lu T, Zhang R Q, *et al.* *Chinese Science Bulletin*, 2000, **45** (15): 1413~ 1417
- Crippen G M, Havel T F. *Distance Geometry and Molecular Conformation*. New York: Chemometrics Research Studies Press, 1988. 101~ 156
- Kuntz I D, Thomason J F, Oshiro C M. *Distance geometry*. *Methods Enzymol*, 1989, **177**: 159~ 204
- Güntert P, Braun W, Wüthrich K. Efficient computation of three-dimensional protein structures in solution from NMR data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J Mol Biol*, 1991, **217** (3): 517~ 530

## A Prediction for The Tertiary Structure of Metallothionein CAP3 From *Colletotrichum gloeosporioides*\*

HE Hong-Zhen, LÜ Tun, ZHAO Nan-Ming, ZHANG Rì-Qing, LIU Jir-Yuan\*\*

(Department of Biological Science and Biotechnology, Tsinghua University,  
State Key Laboratory of Biomembrane and Biotechnology, Beijing 100084, China)

**Abstract** Based on the experimental data of metallothionein-I (MT-I) tertiary structures, interatomic distance constraints of two characteristic structures (CXC, CXXC sequence pattern and metal-Cys chelate tertiary structure) were constructed. Then the distance geometry method was used to work out a number of possible structures, from which those with remarkably lower target function value were selected by a statistical analysis as the predicted tertiary structure models. The predicted structure of blue crab by this method was similar to the experimentally determined structure showed that this method could be used to MT prediction. The tertiary structure of a metallothionein CAP3 from *Colletotrichum gloeosporioides* was modeled. The predicted structure is similar to those of its homologous MT proteins in metal-Cys combination, and reasonable in its structure energy as well as in its fold motif.

**Key words** metallothionein CAP3, tertiary structure prediction, distance geometry method

\* This work was supported by grants from the High-Tech "863" Project for Young Scientists, Trans Century Training Program from the Ministry of Education and State Key Basic Research and Development Plan (G19998010100).

\*\* Corresponding author. Tel: 86-10-62772243, E-mail: liujy@mail.tsinghua.edu.cn

Received: July 30, 2001 Accepted: October 30, 2001